

Garrod, S. and Pickering, M. J. (2016) Dual-stream accounts bridge the gap between monkey audition and human language processing: comment on “Towards a Computational Comparative Neuroprimatology: Framing the language-ready brain” by Michael Arbib. *Physics of Life Reviews*, 16, pp. 69-70. (doi:[10.1016/j.plrev.2016.01.008](https://doi.org/10.1016/j.plrev.2016.01.008))

This is the author’s final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

<http://eprints.gla.ac.uk/120073/>

Deposited on: 14 June 2016

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

Dual-stream accounts bridge the gap between monkey audition and human language processing: Comment on “Towards a computational comparative neuroprimatology: Framing the language-ready brain” by Michael Arbib.

Simon Garrod¹, Martin J Pickering²

1. Institute of Neuroscience and Psychology, University of Glasgow,
2. School of Philosophy, Psychology and Language Sciences, University of Edinburgh

Over the last few years there has been a resurgence of interest in dual-stream *dorsal-ventral* accounts of language processing (4). This has led to recent attempts to bridge the gap between the neurobiology of primate audition and human language processing with the dorsal auditory stream assumed to underlie time-dependent (and syntactic) processing and the ventral to underlie some form of time-independent (and semantic) analysis of the auditory input(3,10). Michael Arbib (1) considers these developments in relation to his earlier Mirror System Hypothesis about the origins of human language processing (11).

Arbib’s review covers two distinct issues: first, how joint action and attention support a dyadic model of communicating brains (leading from complex imitation to pantomimic gesture to proto-language), and, second how a dual-stream account of vision can be related to a dual-stream account of audition and ultimately language processing. Arbib makes many interesting suggestions about the relationship between gesture and language and between the dorsal and ventral streams in vision and audition. But he does not directly relate his analysis of dual-stream processing to the dyadic model of communicating brains developed in section 1.4 of his review. By contrast we suggest that dorsal stream processing (in particular predicting the timing of speech) may contribute to the fluent and rapid turn-taking seen in dialogue. This would relate the dual-stream approach to an analysis of interactive communicating brains.

Arbib takes exception to the claim in (2) that the dorsal auditory stream deals with time-dependent syntactic analysis of the input whereas the ventral stream deals with time-independent semantic analysis, arguing that this is inconsistent with a Template Construction Grammar approach which treats time-critical syntactic analysis as incorporated with semantic analysis of the speech. This argument does depend on specific assumptions about linguistic analysis. In contrast, one argument for time critical processing in the dorsal stream relates to its role in dynamically up-dating predictions about the timing of the speech input. Recent neuroimaging studies indicate that low frequency delta and theta band (1-3Hz; 4-7Hz) neural oscillations in auditory cortex synchronise with delta and theta band oscillations in the speech envelope, thus tracking and predicting speech rate (6, 14). It is assumed that this brain-speech synchronization aids syllabic parsing and prediction of other important temporal events in the speech stream since these oscillations in the speech correspond to syllabic (theta) and other suprasegmental (delta) events (e.g., intonational phrases). Crucially, the degree of synchronization in left auditory cortex is modulated by top-down signals emanating from left frontal and motor areas – a

finding that implicates dorsal stream control of the synchronization and prediction process (9).

Garrod and Pickering (4) propose that such a mechanism may help interlocutors orchestrate turn-taking in dialogue. As Levinson (7) points out, the timing of turn-taking is extremely precise with modal gaps between turns of around 200msecs and overlaps of around 100 msecs. This precise temporal coordination may depend on coupling the timing of speech comprehension and production with neural oscillators under the control of motor cortex (see 13,8).

Turn-taking mechanisms may also shed light on the evolution of interactive communication in primates other than man. As Levinson (7) notes, basic turn-taking systems are found in gestural or vocal communication in prosimians, monkeys and apes. Furthermore, there is evidence that vocal turn-taking even in marmosets relies on coupled oscillators between the 'conversing' monkeys, similar in certain ways to that proposed for human conversational turn-taking (12). Hence understanding the neural basis of turn-taking systems and how it relates to dorsal-ventral auditory streams may well contribute to Arbib's 'computational comparative neuroprimatology'.

Refs:

- (1) Arbib, M. (2016) Towards a Computational Comparative Neuroprimatology: Framing the language-ready brain. *Physics of Life Reviews*. This issue.
- (2) Bornkessel-Schlesewsky, I & Schlewsky, M. (2013) Reconciling, Time, space and function: A new dorsal-ventral stream model of sentence comprehension. *Brain and Language*, 125, 60-76.
- (3) Bornkessel-Schlesewsky, I., Schlewsky, M., Small, SL. & Rauschecker JP Neurobiological roots of language in primate audition: common computational properties, *Trends in cognitive sciences* 19 (3), 142-150
- (4) Hickok, G. & Poeppel, D. (2004) Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92, 67-99.
- (5) Garrod S and Pickering MJ (2015) The use of content and timing to predict turn transitions. *Front. Psychol.* 6:751.
- (6) Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. *PLoS Biology*, 11(12).
- (7) Levinson, S.C. (in press) Turn-taking in human communication – Origins and implications for language processing. *Trends in Cognitive Sciences*.
- (8) O'Dell, M.L., Nieminen, T., and Lennes, M. (2012). Modeling turn-taking rhythms with oscillators. *Linguist. Ural.* 48, 218–227.
- (9) Park, H. InceRR Schyns, P, Thut, G Gross, J. (2015) Frontal Top-Down Signals Increase Coupling of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. *Current Biology*, 25, 1649–1653
- (10) Rauschecker, JP & Scott, SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience* 12, 718 - 724 (2009)
- (11) Rizzolatti, G., Arbib M, A. (1998) Language within our grasp *Trends in Neurosci.* 21(5), 188–94.

- (12) Takahashi, DY, Narayanan,DZ & Ghazanfar1, AA (2013) Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys Current Biology, 23, p2162–2168,
- (13) Wilson, M., and Wilson, T.P. (2005). An oscillator model of the timing of turn-taking. Psychon. Bull. Rev. 12, 957–968.
- (14) Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, et al. (2013) Mechanisms underlying selective neuronal tracking of attended speech at a “cocktail party”. *Neuron* 77: 980–991